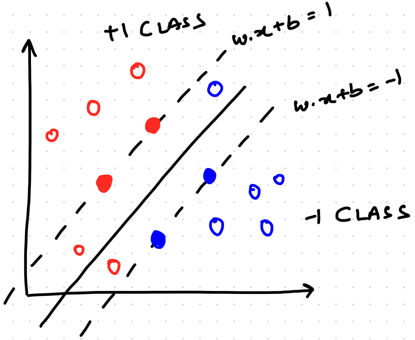
SVM Soft Margin Classification

Nipun Batra

IIT Gandhinagar

July 30, 2025

"SLIGHTLY" NON-SEPARABLE DATA



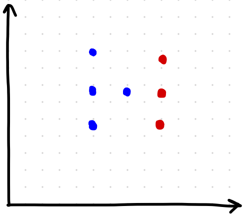
In Dual:

$$\text{minimize} \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$$

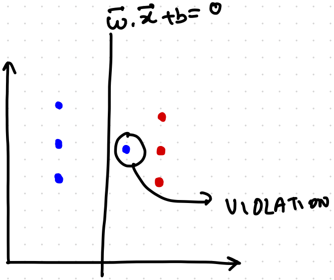
s.t.

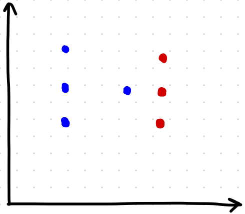
$$0 \leq \alpha_i \leq C \quad \& \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

BIAS-VARIANCE TRADE-OFF



BIAS-VARIANCE TRADE-OFF





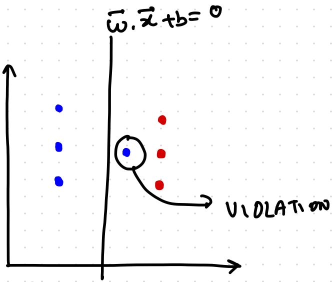
LOW C

LOW PENALTY FOR VIOLATION

HIGH TRAIN ERROR

HIGH BIAS

BIAS-VARIANCE TRADE-OFF



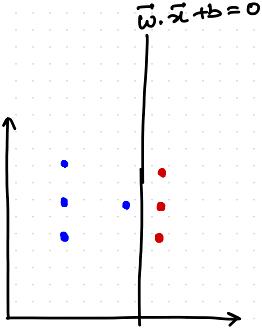
LOW C

LOW PENALTY FOR VIOLATION

HIGH TRAIN ERROR

HIGH BIAS

BIG MARGIN



HIGH C

HIGH PENALTY FOR VIOLATION

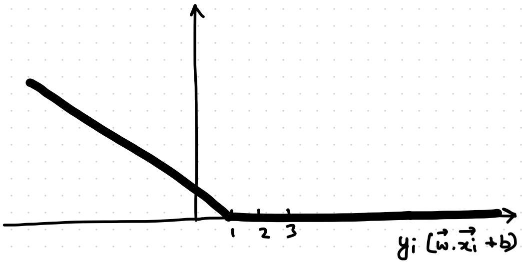
HIGH VARIANCE

SMALL MARGIN

b$\mathrm{w} \cdot \mathbf{x}_i + b)$ 12$\|\mathbf{w}|$

$$\text{minimize} \underbrace{\sum_{i=1}^{N} \max\left[0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\right]}_{\text{Loss}} + \underbrace{\frac{1}{2C}\|\mathbf{w}\|^2}_{\text{Regularisation}}$$

HINGE LOSS



Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\right]$

- Case I $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$
  Lies on Margin: $Loss_i$ = 0

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\right]$

- Case I $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$
  Lies on Margin: $Loss_i$ = 0
- Case II
  $$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1$$
  $Loss_i = 0$

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\right]$

- Case I $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$
  Lies on Margin: $Loss_i$ = 0
- Case II
  $$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1$$
  $Loss_i = 0$

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\right]$

- Case I $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$
  Lies on Margin: $Loss_i$ = 0
- Case II
  $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1$
  $Loss_i = 0$
- Case III
  $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) < 1$
  $Loss_i \neq 0$

Hinge Loss Continued

Q) Is hinge loss convex and differentiable?

Convex: ✓

Differentiable: X

Subgradient: ✓

SVM Loss is Convex

Hinge Loss $\sum(\max[0, (1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b))]$ is convex

Penalty $\frac{1}{2}\|\mathbf{w}\|^2$ is convex

$\therefore$ SVM loss is convex